King's Research Portal

Document Version
Publisher's PDF, also known as Version of record

Link to publication record in King's Research Portal

Citation for published version (APA):
Smithies, J., Westling, C., Sichani, A-M., Mellen, P., & Ciula, A. (2019). Managing 100 Digital Humanities Projects: Digital Scholarship & Archiving in King's Digital Lab. *Digital Humanities Quarterly*, *13*(1), [000411]. http://www.digitalhumanities.org/dhq/vol/13/1/000411/000411.html

Citing this paper
Please note that where the full-text provided on King's Research Portal is the Author Accepted Manuscript or Post-Print version this may differ from the final Published version. If citing, it is advised that you check and use the publisher's definitive version for pagination, volume/issue, and date of publication details. And where the final published version is provided on the Research Portal, if citing you are again advised to check the publisher's website for any subsequent corrections.

General rights
Copyright and moral rights for the publications made accessible in the Research Portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognize and abide by the legal requirements associated with these rights.

•Users may download and print one copy of any publication from the Research Portal for the purpose of private study or research.
•You may not further distribute the material or use it for any profit-making activity or commercial gain
•You may freely distribute the URL identifying the publication in the Research Portal

Take down policy
If you believe that this document breaches copyright please contact librarypure@kcl.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.

Download date: 06. May. 2023

Current Issue
2020: 14.3
Preview Issue
2020: 14.4
Previous Issues
2020: 14.2
2020: 14.1
2019: 13.4
2019: 13.3
2019: 13.2
2019: 13.1
2018: 12.4
2018: 12.3
2018: 12.2
2018: 12.1
2017: 11.4
2017: 11.3
2017: 11.2
2017: 11.1
2016: 10.4
2016: 10.3
2016: 10.2
2016: 10.1
2015: 9.4
2015: 9.3
2015: 9.2
2015: 9.1
2014: 8.4
2014: 8.3
2014: 8.2
2014: 8.1
2013: 7.3
2013: 7.2
2013: 7.1
2012: 6.3
2012: 6.2
2012: 6.1
2011: 5.3
2011: 5.2
2011: 5.1
2010: 4.2
2010: 4.1
2009: 3.4
2009: 3.3
2009: 3.2
2009: 3.1
2008: 2.1
2007: 1.2
2007: 1.1
Indexes
Title
Author

ISSN 1938-4122

Announcements
Call for Reviewers
Call for Submissions

2019
Volume 13 Number 1

2019 13.1 | XML | Discuss (1 Comment)

Managing 100 Digital Humanities Projects: Digital Scholarship & Archiving in King's Digital Lab

James Smithies <james_dot_smithies_at_kcl_dot_ac_dot_uk>, King's College London
Carina Westling <cejwestling_at_gmail_dot_com>, King's College London
Anna-Maria Sichani <amsichani_at_gmail_dot_com>, King's College London
Pam Mellen <pamela_dot_mellen_at_kcl_dot_ac_dot_uk>, King's College London
Arianna Ciula <arianna_dot_ciula_at_kcl_dot_ac_dot_uk>, King's College London

Abstract

During the 2016–2017 financial year, King's Digital Lab (King's College London) undertook an extensive archiving and sustainability project to ensure the ongoing management, security, and sustainability of ~100 digital humanities projects, produced over a twenty-year period. Many of these projects, including seminal publications such as *Aphrodisias in Late Antiquity, Inscriptions of Roman Tripolitania, Henry III Fine Rolls, Jonathan Swift Archive, Jane Austen Manuscripts, The Gascon Rolls, The Gough Map*, and *Inquisitions Post Mortem*, occupy important positions in the history of digital humanities. Of the projects inherited by the lab, about half are either of exceptionally high quality or seminal in other ways but almost all of them struggled with funding and technical issues that threatened their survival. By taking a holistic approach to infrastructure, and software engineering and maintenance, the lab has resolved the majority of the issues and secured the short to medium term future of the projects in its care. This article details the conceptual, procedural, and technical approaches used to achieve that, and offers policy recommendations to prevent repetition of the situation in the future.

1 Digital Humanities (DH) research has reached an inflection point. On the one hand it appears to be in robust health, with an active community spread around the world, well-attended annual conferences, several well-established centres of excellence (be they labs, institutes, or departments), and new initiatives appearing on a regular basis. Activity is particularly strong in the United Kingdom, North America, and Europe, with significant work being done in Asia and Australasia, and important new initiatives developing in South America and Africa. University courses are proliferating at graduate and undergraduate level, and advances are being made to pedagogy [Hirsch 2012]. Intermittent criticism of the field is a sign of increasing intellectual vitality, as colleagues in neighbouring disciplines question its popularity and interrogate its intellectual, ideological, and ethical foundations [Eyers 2013] [Allington et al. 2016]. This activity has appeared at the same time as the notion and utility of DH infrastructure has been questioned [Rockwell 2010], and project teams have been forced to explore ways projects can not only be sustained but elegantly ended #carlin2006. Despite inheriting a relatively deep tradition, we are only beginning to understand the dense entanglements that accrue over time between digital humanists, administrators, and the web servers, programming languages, and tools, we use to produce our publications.

2 While technical digital humanities teams now have a vastly more sophisticated range of options than previous generations, including corporate-grade cloud services and free online services, this has done little to ease the problem of maintainability or sustainability - especially for high quality digital scholarship. Idiosyncratic solutions to specific research questions in this emerging field have left us with a legacy corpus developed from the 1990s into the 2000s, which raises new challenges in terms of sustainability. Problems that have been deferred for years, sometimes decades, have become pressing. A generation of legacy projects that need maintenance but are out of funding have reached critical stages of their lifecycles, an increasingly hostile security context has made DH projects potential attack vectors into institutional networks, heterogeneous and often delicate technologies have complicated the task of maintenance, and an increasing number of emerging formats have made archiving and preservation yet more difficult. This presents a significant, and growing, challenge for the community – and one that needs to be resolved by raising awareness of the issues, evolved management of digital humanities infrastructure, attention to the full lifecycle of projects, and inventive approaches to funding that extend the life and impact of valuable research by sharing costs across funding agencies and institutions. This article aims to contribute to that process by initiating a conversation and explaining the experience and some solutions implemented in King's Digital Lab (KDL) but does not aim to present a straight-forward "How To" guide for other teams. The realisation of robust and holistic approaches to the maintenance of digital research outputs is a matter of some urgency, but no single solution will work for every digital humanities team.

3 It is clear, however, that sustainable management of digital outputs that have survived beyond their initial funding has become a major problem. It is time to admit our problems and share our conceptual and procedural solutions. Such projects, although of central importance to the wider field of digital humanities and humanities scholarship generally, present a range of challenges. In the academic and financial year 2016-2017[1], KDL worked on 6–8 funded projects at any one time and was involved in external grant proposals with a total value of £26m (GBP), together with collaborators across a wide variety of universities and cultural heritage organisations in London, the United Kingdom, Europe, and the United States. This constituted the lab's primary activity and is at the core of its raison d'être. In the same year, however, the lab completed assessment (followed by archiving, migration, or upgrade) of ~100 digital humanities projects undertaken over twenty years of activity at King's College London and inherited from earlier instantiations of DH, including the Centre for Computing in the Humanities (CCH) and the Centre for eResearch in the Humanities (CeRch). Many projects were inherited from the Department of Digital Humanities (DDH), which the lab evolved from and has a close relationship to. This corpus of publications represents valuable and impactful research as well as significant investment from funding bodies, and research and heritage institutions. Humanities scholars rely on and make ongoing reference to the work contained in it, and it is increasingly being integrated into global Linked Open Data initiatives supported by libraries, archives, museums, and other digital humanities teams.

4 Finding a comprehensive and scalable approach to sustainable development in digital humanities labs is a non-trivial problem. Any solution must be tailored to the local environment and help support not only the complexity and range of digital scholarship, but financial and operational issues and more fundamental problems related to entropy of software systems and digital infrastructure. It also needs to allow for the fact that digital tools and infrastructure do not allow for perfect process, perfect archiving, or perfect security: at some point it is always necessary to retreat to principles of risk management and cost-benefit analysis. The work presented here involved coordination with technical specialists, researchers, administrative and financial university staff, and colleagues in IT and the library. The developed process helped KDL transition many digital humanities projects from an insecure to a sustainable basis, but the work is incomplete and will - in a fundamental sense - never end. Some projects, moreover, cannot be "saved" despite best intentions. Rather than aim for perfect process, KDL have chosen to accept archiving and sustainability as a permanent issue that requires ongoing care and attention. It has been added to the lab's Software Development Lifecycle (SDLC) engineering process and is considered from our very first conversations with new project partners. Our experience is shared here to open a conversation and, rather than proposing simple solutions or demanding policy change, to invite discussion.

King's Digital Lab: Background

5 King's Digital Lab (KDL) was launched in November 2015 at King's College London. The lab evolved from the Centre for Computing and the Humanities (established 1995)[2] and the Centre for eResearch in the Humanities (established 2008), which merged to form the Department of Digital Humanities (DDH) in 2012. At the time of writing, DDH delivers 5 masters programmes, 1 PhD programme and an undergraduate programme to ~500 students and comprises ~40 academic staff. KDL was founded to increase digital capability and generate external grant income for digital projects within DDH, and across the Faculty of Arts & Humanities as a whole: it is an independent Arts & Humanities department in its own right, specialising in digital humanities software development but increasingly working with social scientists too. Team members sometimes act as Principal or Co-Investigators on grants but always work in unison with colleagues in other departments and/or institutions, implementing a model for digital humanities research at scale.

6 Unlike core academic departments, which engage in teaching as well as research, KDL is dedicated to research software engineering (RSE), and the implementation of the systems, infrastructure, tools, and processes that are needed to produce digital scholarly outputs. The lab has 12 permanent full-time staff to support these activities: research software analysts, engineers, designers, a systems manager, a project manager, and the director, and maintains its own server and network infrastructure. The team work in close collaboration with the university's IT department and evolving University eResearch team. KDL's research philosophy is evolving: it lies at the intersection of human research and technical systems, exploring and exploiting the creative synergies fostered by this encounter to push the boundaries of digital humanities forward. Taking an active interest in research methodology as well as inevitable business and technical realities, the lab embraces problems we believe are integral to the evolution and sustainability of the field.

7 Conflating the scholarly and cultural and operational aspects of the lab is both an overt act of historicisation - an acknowledgment of the reality of digital scholarship in early 21st century higher education - and a pragmatic response to the inherited and emergent issues outlined in this paper. The design and engineering of software and its supporting infrastructures is a problem that needs to be conceived as at once technical, political, economic, and *human*. While the lab exists to engage in technical development, it is mandated to explore the epistemic and methodological implications of digital humanities development and can contribute to the broader field from a unique vantage point. Its institutional setting, technical expertise, and exposure to research problems that only time can generate positions it to explore fundamental issues of digital theory and method (including but not limited to digital entropy), while at the same time developing innovative methods for new research.

Legacy Portfolio

8 The legacy portfolio supported by KDL is not unique, but significant for its range and scholarly value: it represents a key corpus in the history of digital humanities. Digital Humanities at King's College is indebted to a group of people who were instrumental in developing a range of projects inherited by KDL. Colleagues like Harold Short, John Bradley, Willard McCarty, Charlotte Roueché, Marilyn Deegan, and Paul Spence, were involved in a remarkable array of projects of enormous scholarly value. In collaboration with PIs, both at King's College and in partner institutions, their work provided the core of the lab's inheritance including flagship projects such as *Aphrodisias in Late Antiquity, Inscriptions of Roman Tripolitania, Henry III Fine Rolls, Jonathan Swift Archive, Jane Austen Manuscripts, The Gascon Rolls, The Gough Map, Sharing Ancient Wisdoms, Prosopography of Anglo-Saxon England, Prosopography of the Byzantine World, The Complete Works of Ben Jonson, The Heritage Gazetteer of Cyprus*, and the *Profile of a Doomed Elite*. Work in palaeography by Peter Stokes and Stewart Brookes has prompted a range of projects, including *DigiPal, Models of Authority, Exon Domesday*, and the new *Archetype* framework.[3] This work was delivered in close collaboration with leading technical figures in digital humanities in the United Kingdom, including many who now work in, and with, King's Digital Lab.[4]

9 Of the 100 projects inherited by the lab, about half are either of exceptionally high quality or seminal in other ways. This is a sizeable "estate" to manage, but the authors are aware of at least one team managing considerably more projects, and more than one team who have suffered serious security breaches because of unmaintained applications. Teams struggling with the issues are located in the United States and Europe as well as the United Kingdom, suggesting any issues with policy and approach transcend national borders.[5] Such circumstances entail a considerable moral bind: either ignore the demands of (some) project owners that their projects' digital publications and data continue in perpetuity and turn them off (risking reputational damage and reducing the number of DH projects available to users, more often than not initially supported via public funding), introduce financial risk by maintaining them gratis (absorbing unfunded maintenance costs and undermining other activities), or do nothing and accept the existential risks that accompany a major security breach.

10 Little support is offered from the surrounding culture. Funding agencies might require data management plans to ensure content is gracefully handled, depending on the country of origin, but appear unable to deal with the complex issues associated with the systems that generate and store that data. Collaborators often become uneasy at the use of "industry" frameworks and "business" language, suggesting (understandably) that it detracts from academic research culture. Meanwhile, some critics of the digital humanities appear to be unaware that a universe of very high quality, bespoke, but at-risk digital scholarship exists far away from the values and commercial imperatives of Silicon Valley ideology.[6] In that sense, this paper is an account of a course charted between Charybdis and Scylla, seeking to protect a cargo of scholarship from technical and financial realities, the barbs of critics, the corporatisation of higher education, and gaps in national policy. Were it not for the fact that this is *the precise set of operational tensions* that drives the intellectual and creative culture of laboratories like KDL, and the support of an almost uniformly understanding group of project owners and stakeholders, the combined pressures would be insufferable. Given this, we view this article as an opportunity to articulate the issues facing teams like KDL, gesture towards some of our solutions, and make it easier for other teams to share their experiences and request the resources needed to mitigate issues.

11 The ~100 projects inherited by KDL range across several DH sub-disciplines, with a focus on Digital Classics (23 projects), Digital History (23 projects), and Digital Literary Studies (14 projects). Another group can be best described as Digital Humanities (20 projects), with smaller but important groups in Digital Musicology (5 projects), Cultural Studies (5 projects), and Spanish Studies (4 projects). A further 5 projects are best described as inter-disciplinary. New projects appear on a regular basis, of course, meaning the precise numbers constantly shift. Surprisingly, and accepting that five years is a long time in the digital world, 77 of the projects are less than 5 years old, with only 22 projects more than five years old. Of more concern is the fact that, when KDL was established, the majority of these projects were "orphaned", and left without funding for maintenance. In lieu of merely shutting them down, they had been kept live with little or no maintenance, resulting in some unacceptably old operating systems remaining in production. This is by no means out of step with the situation at many organisations (commercial or otherwise). It reflects an era in the history of computing when technological optimism was somewhat higher and security risks somewhat lower than they are today.

12 We would like to note, in this context, that our openness in publishing the details of the situation is relatively unusual and should indicate the importance we feel the subject holds for the global humanities and social science communities, and the library and archival teams that support them. We have a good degree of confidence the issues have been resolved, as far as is possible given today's environment and the evolving security threats it presents, but – more importantly – feel it is time to have an open conversation about these issues. Teams like KDL struggle with issues presented by myriad pressures: it is neither fair nor productive for Principal Investigators (PIs), funding agencies, and the wider community, to have the reality of those pressures hidden from them. Significantly, rather than seeing such issues as embarrassments, to be hidden from administrators, funders, and colleagues inside and outside our institution, the lab recognises them as research opportunities for developing enhanced methodologies. It can be noted, too, that this attitude represents continuity with the history of digital humanities at King's College rather than a departure from it. Previous generations of colleagues, including Harold Short, Marilyn Deegan, Lorna Hughes, and Sheila Anderson, tried to prompt policy change at a national level (most clearly through the Arts & Humanities Data Service, but also through regular connection with national funding bodies and other organisations), but their efforts were not supported at crucial moments [Rusbridge 2007]. Our goal is to empower similar teams to seek and secure the support needed to do their jobs, and contribute to the development of guidelines, standards and policies that can guide digital scholarship. This ambition needs to be seen in a wider context that includes issues of not only technical and financial sustainability but equitable career paths, ethical attribution, diversity, and DH in developing countries.

13 Although (again) by no means unusual, the details of KDL's technical estate in late 2015 would give many systems administrators sleepless nights. KDL projects were running: Windows 2003 (2 servers); Windows 2008 (9 servers); Debian 4 (13 servers); Debian 5 (32 servers); Debian 6 (33 servers); Debian 7 (10 servers). The preponderance of Linux servers reduced risk significantly, but the age of many of them was enough to be a risk even before the potential impact of weaponised hacking tools on mainstream institutional systems became clear. All servers were backed up, onto older machines that were adequate but not entirely fit for purpose, and a significant security breach could have led to several days' downtime while the systems were restored to their previous best-known state. It was initially difficult to communicate this to some project owners, who were unaware of the need for infrastructure maintenance, and the risks their servers posed. The WannaCry event prompted a marked change, however. PIs who had previously resisted sharing responsibility for their projects' security immediately allowed KDL to turn off servers until emergency patching had been completed or (in the worst cases), both server and application had been rebuilt. The lab was close to taking this action unilaterally, regardless, for the good of everyone involved. The consciousness-raising that accompanied WannaCry, following its impact on the UK National Health Services (NHS) and other key digital infrastructures, made the process considerably easier [Cellan-Jones and Lee 2017].

14 Solid security requires up to date and regularly patched servers, but also up to date and patched application frameworks (the body of code that enables the websites, databases, archives, and digital scholarly editions end-users interact with), which can be equally difficult to maintain. As with the use of Linux, decisions to build using open source tools lowered risk significantly but did not eliminate the need for regular updates. 26 of the oldest projects were built using Java, but 52 were built using the Python-based web framework Django, which has proven to be relatively secure. When coupled with the bespoke XML-based publishing solutions xMod and Kiln (used for digital scholarly editions), security risk and associated costs were almost entirely removed - but these tools could not be used for every project. The most problematic projects in the legacy portfolio were built using PHP-based frameworks such as WordPress and Typo 3, which were promptly removed from KDL servers wherever possible. Exceptions aside, analysis of the lab's application security validated and renewed our focus on a more limited technology stack based on Linux, Python, Django, and associated supporting tools. Other labs might undertake similar analysis and conclude they should focus on a stack *including* Windows and PHP-based tools and *excluding* Linux and frameworks like Django (to better align to their technical history and capabilities): the issue is a matter of systems maintenance and security, not a reflection of the so-called "programming language wars" [Stefik and Hanenberg 2014]. Experimentation with a range of new technologies continues, particularly in emerging frameworks to support augmented and virtual reality, but long-term support is focused on the core tool set.

Policy Context

15 It is not our intention to propose national policy change in this article (either in the UK or other countries), which requires more insight into the complexities of strategy and funding than we possess, but it is important to note that the projects inherited by KDL, and detailed in this paper, were developed using funding that only supported technical development and limited post launch hosting of projects. Limited or no support existed for significant post-funding system maintenance. In that sense, the funders *themselves* signalled that they did not expect (or were not prepared to support) the development of long-term or permanent digital resources: without the goodwill of colleagues and the host institution most of them would have been closed years ago. Their future was often only discussed tangentially, elided in conversations between technical teams and PIs during the development process, in the optimistic hope "something" would happen eventually, and that either the funding agencies would see the value of the scholarly assets being built, or a national solution would be implemented to protect them – or, in the absence of the realisation of such hopes, that hosting institutions would support them gratis in perpetuity. PIs shifted emotional responsibility onto technical teams, and vice-versa: actual contractual responsibility was normally left undefined.

16 With the benefit of hindsight this was unfortunate, but perhaps inevitable given the lack of knowledge about the many intersecting issues in play. Many of the projects hosted by KDL were produced during seminal years in the history of the field, when flagship digital humanities projects demonstrated the potential that digital tools and methods held for arts and humanities research, and they are consequently of considerable cultural and scholarly value. The spirit of 1990s cyber-utopianism - which assumed electronic media would be cheap and technically straightforward to maintain, and that libraries would develop subscription models able to support bespoke non-commercial projects - held back proactive funding of archiving and sustainability initiatives [Turner 2008]. Funding agencies and researchers alike assumed that their role was to prompt expansion and illustrate possibilities, and that issues of maintenance and sustainability would be resolved in the future. This attitude was understandable, but it is having a serious impact on teams who have inherited multiple high profile (and now unfunded) projects that are well beyond their initial funding periods. That is not to suggest that earlier generations of digital humanists did nothing to plan for the future, however. UK colleagues often cite the defunding of the Arts & Humanities Data Service[7] and the AHRC ICT Methods Network[8] at the start of the millennium as signal events that undermined the future of multiple projects.

17 It is reasonable to view this as an international problem. Other UK digital humanities teams report similar issues to KDL, and colleagues inform us that policy gaps have created similar problems in the United States. The problems exist in continental Europe but are less pronounced because of longer-term commitments to infrastructure development and better alignment to STEM-based initiatives that are actively exploring ways to improve Research Data Management (RDM) infrastructure and processes [European Commission 2017] [Rosenthalter et al. 2015]. It is important at the outset to recognise that the issue runs deeper than straightforward problems of IT "service delivery", however. In large part the issues inherited by KDL are the result of a wider conceptual failure, and an inability (or unwillingness) to search "for critical and methodological approaches to digital research in the humanities *grounded in the nature of computing technology and capable of guiding technical development as well as critical and historical analysis*" [Smithies 2017, 3]. If practical work in the digital humanities is to continue, this attitude needs to be fostered, and extended towards the ongoing maintenance, archiving and preservation of projects as well as their development. In an article in *Aeon* in 2017, historians of computing Andrew Russell and Lee Vinsel point out that the technology industry is so ideologically biased towards "newness" that it glosses over the need for maintenance despite it being a significant aspect of the contemporary digital world [Russell and Vinsel 2016]. More pointedly in the context of digital humanities, Paul Edwards et al note that "*sustainable knowledge infrastructures must somehow provide for the long-term preservation and conservation of data, of knowledge, and of practices…*" and that this "*requires not only resource streams, but also conceptual innovation and practical implementation*" [Edwards et al. 2013, 8].

18 The issue has cascading implications for digital humanists and policy makers alike. If digital humanities projects become known for not only soaking up valuable money that could be used in other disciplines, but using that money on unsustainable projects, the central *raison d'être* of the wider tradition - using digital tools and methods to answer research questions in the humanities - will be undermined. However, there is no reason the worst scenarios (permanent loss of multiple flagship digital humanities projects) should come to pass. As Smithies has argued elsewhere, a wider view of digital humanities infrastructure, in its technical as well as intellectual and ethical dimensions, can provide perspectives that aid not only technical development and management, but the development of ethical perspectives, and greater purchase over business decision-making and funding policy [Smithies 2017, 113–151]. Only by exploring this wider perspective can an appropriate understanding be gained, and supporting policy developed. Patrik Svensson takes a similar approach in his recent book about DH infrastructure [Svensson 2016]. That book aligns well to emerging trends in critical infrastructure studies [Liu et al. 2018], platform studies [Montfort and Bogost 2009], maker culture [Sayers 2017], minimal computing [Smithies 2018], and various critical and philosophical approaches perhaps best described as "epistemologies of building" [Ramsay and Rockwell 2012].

19 The problem is that this work tends to be only tangentially related to, or simply ignore, the seemingly pedestrian problems associated with technical design and development, archiving, and sustainability. Work on humanities research infrastructure is often written by people more invested in the easy development of new projects (and thus the easy availability of development teams and server and hosting infrastructure) than their ongoing maintenance, which hinders rather than helps the sustainability argument [Anderson 2013]. The situation is further complicated by widespread cynicism about large-scale infrastructure development resulting from the failure of programmes such as Project Bamboo in the United States [Dombrowski 2014], which aimed to create a large national cyberinfrastructure in the humanities but foundered due to poor requirements elicitation, a focus on service-oriented architecture, and over-use of dominant STEM models. Geoffrey Rockwell is correct to suggest that digital humanities needs to assess its own requirements, and not assume that infrastructures designed for one purpose will fit another, but it is sensible to at least align the digital humanities to approaches in other fields [Rockwell 2010].

20 The work of researchers like Deb Verhoeven and Toby Burrows, who explore the political and aesthetic implications of large-scale Research Infrastructures (RIs) alongside issues of sustainability and maintenance, provides a new model for thinking through these issues [Verhoeven and Burrows 2015] [Verhoeven 2016]. The value of such work issues from its connection of DH infrastructure development and its maintenance with sociological and anthropological work in infrastructure studies capable of normalising technical infrastructure as a *human* and *community* asset in need of maintenance and support, rather than a technical artefact in need of service management [Bowker 1997] [Dourish and Bell 2014]. Well financed infrastructure combined with careful requirements analysis, tailored to the needs of humanities researchers and their local institutions, can dramatically increase the quality (and lower the costs) of digital humanities support, maintenance, and archiving, but lack of technical leadership has stymied development. Effort also needs to be directed towards the development of best practice and quality assessment frameworks for digital scholarship that include sustainability and maintenance at their core.

21 These perspectives are informed by changes in the policies guiding the development and management of STEM RIs which, although larger in scale, deal with many of the same issues and are not as focused on technology as sceptics might assume. A 2017 European Commission working paper on sustainable research infrastructures noted the centrality of both people and technology to the future of reproducible science [European Commission 2017], and a number of reports on e-infrastructure at the European level and in the United Kingdom have made similar recommendations [ESF 2011] [Ciula, Nyhan, and Moulin 2013] [Open Research Data Task Force 2017]. The 2017 "State of the Nation" report of the UK Research Software Engineering (RSE) association overtly positions permanent career paths at the core of both high-quality science, and technical sustainability [Alys et al. 2017]. If there is a failure of post-millennium digital humanities, it could well be related to this human aspect, rather than anything overtly technical: setting aside all other considerations, permanent DH development teams will resolve most issues of sustainability and maintenance.

22 The experience of KDL suggests that the most effective strategy is to offer open-ended contracts and then embed archiving and maintenance deep into the culture of technical development, from requirements definition and the identification of digital research tools and methods, through to infrastructure design, deployment and maintenance. This is based on a conception of infrastructure that moves beyond material technical necessities, templates, and process documents (as essential as they are), towards one that acknowledges the centrality of people, funding, ethics, technology strategy, software engineering method, and data management to the long-term health of our research infrastructures. This becomes even more pressing if we acknowledge the wider epistemological and methodological shifts occurring across scientific and humanistic disciplines, related to the emergence of data science but also myriad new forms of research dissemination and product development. The community needs to recognise that high quality research requires attention to long-term digital sustainability if quality is to be maintained. This extends well beyond the specifically *digital* humanities, of course, and relates to all disciplines and interdisciplinary efforts that use digital tools and methods. Importantly, the failure (or sub-optimal performance) of previous large-scale infrastructure efforts supports the argument for greater attention to the need for investment in human capital and process maturity alongside capital investment. This suggests the need for a range of initiatives from institutions engaging in DH activity and funding agencies supporting it, from the development of viable technical career paths, to training in basic software development methods: archiving and sustainability is only one aspect.

Software Development Life-cycle (SDLC) & Infrastructure

23 A key part of KDL's work concerns improvements to the engineering and procedural frameworks that enable digital scholarship. Much like research, software development rarely takes a linear path, and the relative volatility of the open web and rapid development of new technologies presents an ecosystem within which published work needs to be protected and maintained over time. Rather than presenting a pristine environment for artefacts, the digital environment, much like the physical one, presents challenges of an economic, political and entropic nature. The precarious existence of artefacts in the physical world, and the evolving responses from the research community to their preservation and documentation, therefore inform our digital practices. To this end, KDL uses an approach to the funding and management of research projects that considers the complexities of not just the research, but also software development and its ongoing sustainability in a changing digital landscape. While slightly increasing initial costs, the benefits of this approach accrue over time - particularly in relation to academic impact, but also medium and long-term maintenance, archiving and preservation.

24 To support this, the team have added System, Application, and Data Lifecycle Management to our Software Development Lifecycle (SDLC), along with Research Data Management. This has resulted in a process of analysis, development, and maintenance underpinned by Service Level Agreements (SLAs) defined in collaboration with PIs and management. The SDLC is based on the Agile DSDM® method [Agile Business Consortium 2016], adapted for a research context. A range of archival products (static sites, removal of front-end, data migration, graceful shutdown, visualisation etc.) are now considered at the initial requirements gathering phase of projects, for implementation when funding ends. Although our concern here is with archiving, maintenance, and sustainability processes for the projects themselves, the work functions within a wider context of not only ongoing research activity but software engineering process and infrastructure management.[9]

25 The laboratory inherited significant infrastructure from the Department of Digital Humanities (DDH): rack servers supporting 400GB RAM, over 180 virtual machines, 27TB of data, and over 100 digital projects ranging from simple WordPress and Omeka sites to ground-breaking scholarly editions and historical prosopographies. At the time of writing a full infrastructure upgrade has been completed, including the deployment of new enterprise backup servers and core infrastructure that has upgraded capacity to ~1TB of RAM and additional disk space running on Solid State Drives (SSD). Network capacity has been upgraded from 1GB to 10GB. The new infrastructure has capacity for significantly more than 200 virtual machines, and planning has already started for a renewal cycle starting in 2023, to ensure continuity past the life of even the new infrastructure. This information is provided less as an advertisement for KDL, than as a reminder that sustainability requires maintenance of supporting hardware as well as the software that is the focus of this article. Coordinating maintenance of all levels of the technology stack requires considerable effort when it needs to support more than a handful of projects.

Principles

26 Our experience suggests that, much as with traditional production and publication of research materials of archival quality, digital projects benefit from being planned and executed with their longevity in mind from the start. This often involves updating scholarly content, but always involves technical maintenance to ensure the publication remains accessible. This places additional importance on consistency and transparency in approach, supported by effective dissemination and internal peer review of technical documentation. Maintenance and ongoing hosting of a digitally published research project needs to be included in grant application budgets, reflecting the life-cycle of the project beyond the date of publication, with a set of maintenance milestones determined at the outset. This is not merely good operational practice, but an indication of the intellectual maturity of the project. Proper understanding of digital scholarship requires an acknowledgement of its entropic nature; the absence of forward planning implies a misunderstanding of the object being produced at a fundamental – perhaps ontological – level. KDL's process is thus guided by a desire to enable high quality digital scholarship, balancing technical and financial issues with the intellectual and historical significance of the project alongside a consideration of its impact, future funding potential, and potential contribution to the UK Research Excellence Framework (REF). The process blends experience with common sense, rather than being anything particularly complicated. It also assumes, significantly, that not all projects should be maintained in perpetuity. Some are better conceived as short-term or even momentary interventions in the scholarly conversation, to be archived online for the historical record but not worth the intellectual, technical, and financial overhead of ongoing maintenance. Convincing PIs of this can sometimes be difficult but the more they consider the wider epistemological import of their work (and often more importantly, the methodological and even ontological purpose of the output they aim to produce) the more open they become. The realisation they don't need to bind themselves to a project permanently - forever concerned about its future maintenance - usually comes as something of a relief.

27 KDL's archiving process (see https://www.kdl.kcl.ac.uk/our-work/archiving-sustainability/) has been, and is still being, developed in response to emergent tensions between the envisioned and manifest material, financial, and political conditions in which legacy projects exist. As Paul Conway has noted, transforming archiving and preservation practice entails fundamental decisions about how the practice is "*conceived, organised, and funded*" [Conway 2010, 69]. Our findings will ideally contribute to a conversation across the digital humanities community, funding agencies, and policy makers with a view to identifying and implementing (or at the least recommending) frameworks, infrastructures, and funding mechanisms that can ensure the sustainability of digital projects and their data in a way that shares the burden between universities and cultural heritage organisations, and funding agencies. While it is unreasonable to expect funding agencies to provide ongoing funding for all projects, it does seem reasonable to ask their support for projects that are managed according to transparent processes and accepted frameworks, that include a range of archival approaches, and integration into Research Data Management (RDM) systems that leading research agencies advocate greater use of [Open Research Data Task Force 2017]. As indicated earlier in this article, we do not feel it is our place to provide detailed recommendations here, however: the issue needs ongoing dialogue and careful consideration across the community.

28 By developing open approaches to the development of archiving and sustainability frameworks, even if they are merely the "*least bad*" option [Conway 2010, 72], the digital humanities community might aspire to deliver on the promise of earlier initiatives like the Arts & Humanities Data Service and safeguard the future of both the community and public investment in digital research projects. It is worth noting here that additional technical work has been initiated behind the scenes at King's College, in a self-funded collaboration between the lab and the DH department. The goal is to create a "data lake" of metadata and digital objects collected over the history of DH at King's, and comprising over 5 million digital objects, for use in teaching as well as research (see https://data.kdl.kcl.ac.uk/). This is part of a commitment to the implementation of deep infrastructure to support DH archiving, which will be aligned to institutional research data management infrastructure and made openly available to the wider community. Work is progressing slowly, as time and funding allows, but the goal is to create a suite of approaches that can be used in the future.

29 Digital curation has been described as a "*new discipline*" [Adams 2009], evolving from archives and libraries tasked with assessing digital material for collection, use, and preservation. KDL's process for archiving inherited "legacy" projects reflects this. Rather than relying on rigid assessment matrices or requiring slavish attention to cost-benefit analyses, it is self-consciously oriented towards relatively subjective issues of "scholarly and intellectual value" and "cultural heritage value". These need to be balanced against hard operational and financial realities, but it was decided relatively early in the process that it would not be possible to create a procrustean assessment framework that could be applied rigorously to all projects: the heterogeneous nature of the projects (technically as well as intellectually), the frequent mismatch between scholarly value and straight-forward impact metrics such as web traffic, as well as uneven access to funding meant that a more holistic - but still consistent and transparent - process needed to be adopted.

30 KDL analysts therefore assess each project in terms of scholarly value, technical complexity, security risk, maintenance cost, infrastructure cost, PI engagement, institutional support, and value to KDL, and value to King's College London. Early assessments used a tabular matrix to guide analysis, but this was quickly abandoned as too limiting: recommendations are made in prose form, allowing quantitative and qualitative issues to be taken into account. A brief "business case", including recommendations and costs, is then presented to the Vice Dean Research, Faculty of Arts & Humanities, and a decision is made. Problematic cases can be referred to the Faculty Research and Impact Team (FRIT), and upwards to the Dean if necessary. The process, at this high level, works very well. Simon Tanner's notion of a "balanced view", assessing value using both subjective and objective measures, allows the lab to act as liaison between the projects and University, and thus support the projects and the wider DH community [Tanner 2012]. The key principle is that KDL acts as facilitator rather than decision-maker, providing professional digital humanities analysis to both PI and management. This requires resources to engage in due diligence, willingness to steward sometimes difficult conversations, and occasional recommendations that projects be archived rather than maintained in their live state, but the process ensures all stakeholders have equal access to information and that escalation paths exist.

Implementation

31 It is worth detailing the effort required to work through KDL's archiving and sustainability issues. During the financial year 2016/2017, the lab undertook a complete audit of all projects held on its servers, including inherited legacy projects, and developed processes for realistic costing of their maintenance and hosting. In tandem, the lab set up contractual agreements that supported the reintegration of the updated legacy projects brought under Service Level Agreements (SLA) into the broader production processes of the lab. During the final four months of the financial year 2016/2017, one full-time member of staff was dedicated to the implementation of the new processes, with the intention to bring all prior legacy projects into current processes under SLA, migrate projects not suited to further managed hosting at KDL to the university's IT department (ITS), external hosting, or a static legacy server, and archiving the remainder. A pilot phase was conducted using the portfolios of two prolific King's College London researchers. Business cases for those projects were submitted to the Faculty of Arts & Humanities, resulting in approval for 5 years' support and maintenance.

32 It quickly became apparent that lengthy documents could be replaced with straightforward, fully itemised and costed Service Level Agreements (SLAs),[10] to clarify the extent and duration of KDL's commitment. These are now issued as part of the release process of any project approaching finalisation and launch, and discussed with PIs in the earliest stages of project definition. As the pilot phase progressed, technical and supporting data about additional projects was gathered, including historical information about funding, PIs/Co-Is, external stakeholders, and infrastructure. This required the identification and synthesis of multiple historical sources but enabled KDL to gain an overview of the extent of the legacy projects, including dependencies and risks. The information was collated and included in documentation that supported the reintegration of each project into the lab's active production cycle, whether that be via managed decommissioning, migration, or defined support and maintenance underwritten by key stakeholders. Based on this high-level assessment, 29 projects were dealt with almost immediately in a first phase that involved them being taken offline and archived by storing database dumps and content files in zip files, because they were incomplete, or incurred security risk out of all proportion to their scholarly value.[11] Others required only basic maintenance to make them secure. A further 35 were scheduled for Phase 2,[12] and 35 for Phase 3.[13] Only legacy projects that were no longer in active development were considered. Another class of project, inherited from DDH but still in active development, were dealt with using a different process. A second key document - the Statement of Work (SoW) - evolved to fit a subsidiary need: to detail and cost work required to bring projects up to an acceptable standard for ongoing hosting. That might only involve simple server upgrades, requiring half a day, or several weeks of active development to rebuild the site in its entirety.[14]

33 King's College London Faculty of Arts & Humanities approved all the business cases presented to it for support of ongoing maintenance and hosting of projects led by Principal Investigators at King's College. It should be remembered that the approved SLAs are all finite - ranging from two to five years - but equally important to note that agreement was reached only after robust business cases were produced, detailing the scholarly and cultural heritage value, the significance to the Research Excellence Framework (REF), the "brand" value to the university, and the PI's future career. This created a new, and essential, level of clarity and made the value of the projects more apparent. The process, which made cost, value, and mutual expectations transparent, was a necessary first step towards the faculty managing its digital assets in a more transparent and cost-effective way, and in alignment to its wider strategic direction. It is perhaps not an ideal solution, which would involve limitless funds and assurances of perpetual support, but it is practical and (we think) sensible given the complexities of long-term technology management and the need to accept competing needs for finite funds.

34 If some projects are eventually moved towards archiving, a decommissioning process is followed, aligned with wider university research data management requirements. It is highly unlikely now that any projects will simply disappear. At the very least their data and a public metadata record will be retained: the future of each project can be discussed on a case-by-case basis. Enhanced transparency has also facilitated co-funding arrangements (between College, Faculty, Department, external partner, and funder, for example), reducing the average SLA cost of ~£2000 GBP per year to an extremely reasonable level for each party. It is equally important to note that KDL is only formally authorised to charge maintenance and hosting at cost recovery level, far below commercial rates (this is the case for normal project work too). This might need to be adjusted in future years, to manage demand if nothing else, but was a crucial element in explaining and justifying the archiving and sustainability projects to colleagues.

35 The process has made us keenly aware of gaps in contemporary funding models, which would ideally incentivise projects to manage their future according to similarly transparent and flexible models, but instead incentivise researchers to produce "orphan" projects with uncertain futures. If a tone of frustration is detected in this article it stems from the relatively common-sense nature of the solutions, coupled with the significant stress placed on teams like KDL by a lack of robust policy. This is not to criticise funding agencies, who have been learning about the implications of digital scholarship alongside the communities they serve (they do an excellent job, with limited resources) but it is important to recognise the human cost of poorly managed projects and infrastructure. It is concerning that recent updates to the UK Arts & Humanities Research Council (AHRC) grant application process is likely to worsen rather than improve the situation in that country, by requiring data management plans but nothing related to system quality, infrastructure, or lifecycle management.

36 Conversations with PIs outside King's College London were often the most difficult, as is to be expected given differences in administrative alignment and awareness of KDL as a new initiative. Expectations of ongoing hosting and maintenance were often ill-defined, and reliant on memory rather than crisp documentation: a result, again, of the loose requirements for archiving and sustainability in past years, as well as changes in personnel and restructuring. In many cases, the production of a SLA was all that was required for the PI to request support from their university (so that they had a simple document to present to administrative teams, usually with only a modest cost attached). If it could be demonstrated that a King's staff member was closely involved in the project or stood to benefit from its ongoing maintenance, King's College London would support a proportion of the SLA. Discussions could become difficult in more complex cases, such as when significant work needed to be undertaken to upgrade the project, or maintenance costs were above the average (normally due to significant use of disk space) but all PIs, internal and external, were offered three scenarios:

1. Service Level Agreements, and (where appropriate) software updates, which guaranteed hosting, regular software maintenance, and server updates under renewable two to five-year contracts, costed on the basis of individual project requirements and including Statements of Work (SoWs), when required, for necessary additional upgrade work.
2. For non-King's staff, migration to the partner institution for local hosting.
3. Archiving of websites no longer in active use. This option did not result in the destruction of research data and could entail rendering websites static for migration to a legacy server, or packaging for archival storage.

37 The last option can present problems, given the complexity of some of the projects and the state of the art in digital archiving. Technical issues abound. A range of "archival solutions" have been considered, ranging from removing complex front-end websites and archiving data, to software emulation, and packaging sites as virtual machines for offline use. The basic philosophy is to embrace heterogeneity of archival solutions, in line with the heterogeneity of the projects themselves. Bespoke approaches are developed on a case-by-case basis, although always in alignment with wider university, national and (where appropriate) international infrastructure initiatives.

38 Other initiatives are being considered too. At the time of writing, KDL is discussing an arrangement with the British Library National Web Archive to improve technical and procedural alignment. King's Research Data Management system is likely to be used for preservation of raw research data along with the lab's own server infrastructure. A project has been completed with the British Museum to produce static sites (more conducive to future archiving) from one of their legacy projects [Jakeman 2018], and a collaboration with Stanford University Press is exploring new modes of digital publishing to balance advanced features with sustainability and maintainability [Ciula 2017]. The internal project referred to above, with the Department of Digital Humanities, aims to aggregate Digital Humanities content stored at King's College and making it publicly available for reuse so that even if some projects do lose their active web presence, their data will still be accessible. The lab is beginning to consider in some technical detail the different options available for archiving and preservation, including the difference between presentation and data layers, the possibility of preserving functionally limited but usable "static" websites rather than complete systems, the possibility of packaging publications into downloadable "virtual machines" that can be run on the desktop, and coupling of all of these approaches with "snapshots" stored in the British Library National Web Archive and Internet Archive. The work described in this paper only becomes tractable to a range of solutions, in other words, conducted using a research-oriented frame of mind that seeks to embed archiving and preservation deeply within core digital humanities theory, method, practice, and policy. Improved policy and infrastructure at a national level would help significantly, but this is a multi-faceted issue that will require broad-based input and support.

39 When it was clear the best possible approach to assessment had been found, transparent processes were in place, and clear options determined, emails were sent to PIs *en masse* to accelerate phases Two and Three. It had become essential the assessment process not drag on, undermining the future of the lab, so there was a degree of nervousness about potential responses. In the initial email to project partners, a deadline for responses within 6 weeks from the sending date was given, after which Faculty would be notified of the status of the resource. After a further month, the permission would be sought from Faculty to archive the projects of non-responsive project partners. This timeline was clearly set out in the emails and followed to the letter. Responses were largely swift and positive, allowing mutually acceptable solutions to be identified in collaboration. Project partners generally responded to initial contact well within the stated time, and often immediately. Responses were broadly appreciative, and the rationale for putting older digital research outputs on secure footing appeared intuitively clear. This raises the question of whether resistance to adopting best practice across the wider research community is exaggerated: it is perhaps more the case that robust methods and clear processes are lacking, and funding policy acts against their development.

40 The lab's attitude, enabled by decisions made within Faculty, prompted progress. King's Digital Lab operates on a non-profit basis (with accordingly slim margins), so one of the most fundamental stages in assessment of the legacy projects was the audit of not just the digital resources held on KDL servers, but also defining the costs involved in their responsible ongoing management and hosting. In this sense the lab performed an administrative and communicative role, rather than acting as judge and jury. The inherently positive nature of the process made it more likely PIs would respond well and allowed the lab to streamline the further processing of legacy projects, and minimise detailed negotiation and problem solving for which there is limited resource. The aim was to conclude the financial year of 2016/2017 with no undocumented or out of contract legacy projects remaining on KDL servers, and all legacy projects that were neither migrated nor archived being brought under Service Level Agreements.

41 That was not completely achieved, but results were satisfactory. At the time of writing all assessment and decisionmaking has been completed, Service Level Agreements are in place for projects that are to remain hosted on KDL servers, migration has occurred or is scheduled for other projects, and archiving of the remainder will occur when time and resource allows. Given no perfect final state will ever be reached the initial task of rationalising and safeguarding the lab's project inheritance can be considered to be complete. The newly established processes will be used to manage the lab's project estate for the foreseeable future. Security risk has been brought within significantly more acceptable tolerances. At the end of the process dozens of once uncertain projects will have been given clarity, and a valuable corpus of digital humanities projects will have been brought under robust management. Surprisingly, given the anxiety that attended the start of the initiative, 46% of the projects were placed under Service Level Agreements, guaranteeing between three and five years of secure maintenance and hosting. Where the end date of projects passed less than five years ago, the lab issued backdated, zero-cost Service Level Agreements, itemised with future costs for each component, to clearly signpost future hosting and maintenance needs. This effectively gave several major projects no-cost extensions to their hosting and maintenance, as well as giving them time to consider their options and plan for the future. At least two significant sites will be rebuilt using new funding, and several others will be subject to follow-on funding proposals. It is worth mentioning that a very small subset of projects (five in total) await full resolution, while discussions around creative funding (e.g. crowdfunding and archiving options) continue.

42 In conjunction with the upgrade to KDL's core infrastructure, this gives our community 3-5 years to continue seeking new options and align to evolving archiving and preservation efforts in the wider research data management and eResearch communities [Nicholson 2018]. 39% of the inherited projects have been archived in some form, 13% on a static HTML legacy server that allows their basic content to remain live but incurs no further maintenance, and 26% on local backup servers. No data, in the form of image files or otherwise, has been removed from potential circulation. Plans are in place to migrate the remaining 15% of the projects to other institutions, in a very pleasing move that signals that they also see the value in investing in the future of digital scholarship. 6% will be migrated to a WordPress service hosted in King's College IT department, and 9% will be migrated to external hosting providers. The onward cost of our current project archiving services are negligible; local backup is supported from baseline operating costs, and the running cost of the two static HTML legacy servers is ~£600 per annum. The major costs, naturally, stem from the 12 months of effort, including 4 months with a dedicated full-time team member, to undertake assessment, produce documentation, and communicate with PIs. It is possible that significant additional costs will appear when more complex sites need to be archived, too, but these cases will appear in a staggered way and therefore be more manageable as part of the lab's normal software development and maintenance process. The end result, in simple terms, is KDL's new "maintenance schedule": a list of ~50 projects, all covered by Service Level Agreements and generating modest internal and external income to offset costs. Concerns remain about some projects, and others remain "in process", but that – in our estimation – is the best that can be expected: maintenance and archiving of digital scholarship is an iterative, continuous process, that does not allow for perfect endings.

Conclusion

43 King's Digital Lab has implemented pragmatic processes that take into account the human, as well as the technical, financial and political perspectives implicit in digital scholarship. It has reinforced the lab's commitment to producing digital research within a holistic and scalable framework, supported by straightforward documentation to ensure mutual clarity about what can be expected from research partnerships. A key component of this framework includes the enhanced Software Development Lifecycle (SDLC) process, which is now implemented from the inception of a project, to align its development with post-publication maintenance and, where appropriate, archiving plans. Early clarity about the feasibility and cost of maintaining projects beyond the funded period allows all parties time to plan ahead, with sufficient time to accommodate the development and turnaround time of follow-on funding applications, negotiations with partner institutions, infrastructure resourcing and requisite allocation of staff time. In addition to optimising maintenance and management of legacy digital research outputs, this approach minimises ambiguity regarding responsibilities and expectations, and contributes to reputation risk management in more than one dimension. Contrary to what might have been expected, KDL's experience of introducing the level of transparency and process described in this article was almost uniformly positive.

44 Successive generations of software (to support visualisation, AR/VR) and other efforts to enhance research methodologies and impact mean the urgency of the questions addressed in this paper is unlikely to diminish in future years, requiring ongoing interrogation of what is an "ideal" technology stack, and best practice. While experimentation with new technology is vital and the precise details of future process design cannot be rigidly determined, more attention to its sustainability, particularly where there is significant investment from public funds, will enhance the field, enhance the benefits of interdisciplinary collaboration outside of the Arts and Humanities, and strengthen arguments for robust funding of the digital humanities sector. Here, it is necessary to differentiate between established technologies and experimental ones. We need to accept brittle code and the possibility of failure in the shorter term for developing technologies but incubate emerging technologies within a context of "legacy risk assessment" informed by industry standards and including upfront analyses of wider infrastructures and technological limitations. Software sustainability will, in all likelihood, remain a pressing issue for the foreseeable future across all research disciplines. The broader conclusion from the experience of KDL is that entropic factors should be taken into account at early planning stages and be accepted by all parties to the project including PIs, developers, and funders. Here, the degree of orientation towards (or away from) archiving and sustainability are core concerns. Funding and associated policy is central to sustainable development, maintenance and archiving. Assuming that future technologies will make it easier or cheaper to solve problems associated with digital entropy is no longer adequate. Sustainable funding strategies need to be based on transparent costing that includes infrastructure and maintenance costs and made simpler and more reliable by established best practice. For this to be effective, realistic costing methods need to be developed and shared between product partners, and embedded within funding policy.

Acknowledgements

45 King's Digital Lab is indebted to the many PIs who initiated and led the projects described in this paper. Special thanks to Harold Short and colleagues (too many to list) in the Centre for Computing and the Humanities (CCH), Centre for eResearch (CERCH), and Department of Digital Humanities (DDH) for many years of ground-breaking work. We would also like to thank Paul Readman, and King's College London Faculty of Arts & Humanities, for their generous support and wise counsel in establishing King's Digital Lab. Paul Readman, Tobias Blanke, Charlotte Roueché, Paul Spence, Simon Tanner, and anonymous *DHQ* reviewers provided valuable feedback on drafts of the article.

Notes

[1] The financial year runs from August to July at our institution.

[2] CCH itself evolved from an initiative known as the Research Unit in Humanities Computing, established in 1992. For more information see [Short et al. 2012]

[3] Projects URLs with associated Principal Investigator (PI) acting as current signatory of the Service Level Agreement (SLA) with KDL: Charlotte Roueché (PI), *Aphrodisias in Late Antiquity* (2005) <http://insaph.kcl.ac.uk/ala2004/index.html>; Charlotte Roueché (PI), *The Inscriptions of Roman Tripolitania* (2009) <http://inslib.kcl.ac.uk/irt2009/index.html>; David Carpenter (PI), *The Henry III Fine Rolls Project* (2009), <http://www.finerollshenry3.org.uk>; James McLaverty (PI), *Jonathan Swift Archive* (2009) <http://jonathanswiftarchive.org.uk/index.html>; Kathryn Sutherland (PI), *Jane Austen's Fiction Manuscripts Digital Edition* (2010) <http://www.janeausten.ac.uk/index.html>; Anne Curry (PI), *The Gascon Rolls project 1317-1468* (2011), http://www.gasconrolls.org/en; Keith Lilley (PI), *Linguistic Geographies: The Gough Map of Great Britain* (2011), <http://www.goughmap.org>; Michael Hicks (PI), *Mapping the Medieval Countryside Places, People, and Properties* (2012), <http://www.inquisitionspostmortem.ac.uk>; Charlotte Roueché (PI) *Sharing Ancient Wisdoms* (2013), <http://www.ancientwisdoms.ac.uk>; John Martindale (PI), *Prosopography of the Byzantine Empire* (2014), <http://www.pbe.kcl.ac.uk>; Martin Butler, *The Cambridge Edition of the Works of Ben Jonson Online* (2014), <https://universitypublishingonline.org/cambridge/benjonson>; Charlotte Roueché (PI), *Heritage Gazetteer for Cyprus* (2015), <http://www.cyprusgazetteer.org>; Stephen Baxter (PI), *Prosopography of Anglo-Saxon England* (2010), <http://www.pase.ac.uk>; Dauvit Broun (PI), *Models of Authority: Scottish Charters and the Emergence of Government* (2017), <https://www.modelsofauthority.ac.uk>; Julia Crick, *The Conquerors' Commissioners: Unlocking the Domesday Survey of SW England* (2017), <https://www.exondomesday.ac.uk>.

[4] At the risk of forgetting important contributors we would like to acknowledge Miguel Vieira, Elliott Hall, Jamie Norrish, Paul Caton, Geoffroy Noel, Gabby Bodard, Arianna Ciula, Neil Jakeman, Charlotte Tupman, Ginestra Ferraro, Elena Pierazzo, Simona Stoyanova, Paul Vetch, Tamara Lopez, Gerhard Brey (†), Raffaele Viglianti, Valeria Vitale, Eleonora Litta, Zaneta Au, Hafed Walda, Richard Palmer, Alejandro Giacometti, Michele Pasin, Faith Lawrence, Peter Rose, John Lee, Jasmine Kelly, Artemis Papakostoulis, Caroline Bearron, Osman Hankir, Felix Herrman, Martin Jessop, Tim Watts, Brian Maher, Andrew Wareham, Juan Garces.

[5] We have chosen not to name them in this article, out of respect for their situation.

[6] See Allington, Daniel, Brouillette, Sarah and Golumbia, David, "Neoliberal Tools (and Archives): A Political History of Digital Humanities." *Los Angeles Review of Books*, May 1, 2016 for an example of the neoliberal critique of digital humanities, and Smithies, James, *The Digital Humanities and the Digital Modern*. Basingstoke: Palgrave Macmillan, 2017 for a rejoinder.

[7] Wikipedia contributors, "Arts and Humanities Data Service," *Wikipedia, The Free Encyclopedia*, https://en.wikipedia.org/w/index.php?title=Arts_and_Humanities_Data_Service&oldid=820259813 (accessed March 1, 2018).

[8] Arts & Humanities Research Council, "AHRC ICT Methods Network: Supporting the Digital Arts and Humanities", 2005–2008. http://www.methodsnetwork.ac.uk/index.html.

[9] Future articles are planned to detail our SDLC, which is too involved to describe in detail here. It is based on the Agile DSDM® method for those interested in exploring more. See https://www.kdl.kcl.ac.uk/how-we-work/why-work-us/ for a high-level overview.

[10] KDL's SDLC templates are made progressively available in our GitHub repository: https://github.com/kingsdigitallab/sdlc-for-rse.

[11] For phase 1, projects were selected on the basis of low research value and minimal complexity, e.g. underused blog sites, orphaned pilot sites, low relevance to King's College and King's Digital Lab (e.g. hosting of personal Wordpress sites, conference sites without updates since the main event, etc). A template email, offering four options (Service Level Agreement, migration to own host, migration to ITS microsite (if a King's partner), or archiving) was sent to projects partners. Recipients were asked to respond within six weeks, after which we said that permission would be sought from Faculty to decommission the web resource. A further month from this, the site would be decommissioned. The majority of project partners responded in a timely fashion, and we successfully agreed on a future path for individual resources.

[12] Phase 2, analysis of the remaining legacy projects (more complex digital research outputs, primarily REF-able and perceived to be of mid- to high research value) began with an investigation within the lab to unearth institutional memory of the affected projects. They were prioritised according to the categories "SLA supported" (where no moneys would be charged to the project in the contractual period), "SLA paid" (charged according to cost recovery including overheads for services rendered over the contractual period), "migration" (to static legacy server maintained gratis by KDL at low cost, or to another host), or "archiving" (with AWS Glacier or similar). Responses were requested within 6 weeks.

[13] In the final phase, we aggregated preferred outcomes from the pilot, first and second phases.

[14] Although we weren't aware of it at the time, this approach aligns well to Matthew Addis' notion of "Minimal Viable Preservation", which recommends an "engineering approach" to preservation that takes care of straight-forward issues before moving onto more complex (and therefore costly) cases. Matthew Addis. "Minimum Viable Preservation - Digital Preservation Coalition." Digital Preservation Coalition, November 12, 2018. https://dpconline.org/blog/minimum-viable-preservation (accessed November 28, 2018).

Works Cited

Adams 2009 Adams, Wright R. "Archiving Digital Materials: An Overview of the Issues." *Journal of Interlibrary Loan, Document Delivery & Electronic Reserve* 19, no. 4 (2009): 325–335.

Agile Business Consortium 2016 Agile Business Consortium, "The DSDM Agile Project Framework (2014 Onwards)", Agile Business Consortium, February 4, 2016. https://www.agilebusiness.org/resources/dsdm-handbooks/the-dsdm-agile-project-framework-2014-onwards (accessed February 21, 2018).

Allington et al. 2016 Allington, Daniel, Brouillette, Sarah and Golumbia, David, "Neoliberal Tools (and Archives): A Political History of Digital Humanities." *Los Angeles Review of Books*, May 1, 2016.

Alys et al. 2017 Alys, Brett, Croucher, Michael, Haines, Robert, Hettrick, Simon, Hetherington, James, Stillwell, Mark and Wyatt, Claire, "Research Software Engineers: State of the Nation Report 2017." Zenodo, April 6, 2017. https://doi.org/10.5281/zenodo.495360 (accessed Febuary 21, 2018).

Anderson 2013 Anderson, Sheila, "What are Research Infrastructures?", *International Journal of Humanities and Arts Computing*, Volume 7 Issue 1-2, pp.4-23, (2013).

Bowker 1997 Bowker, Geoffrey C., *Social Science, Technical Systems, and Cooperative Work: Beyond the Great Divide*. Lawrence Erlbaum Associates, 1997.

Carlin et al. 2016 Carlin, Claire, Czaykowska-Higgins, Ewa, Jenstad, Janelle, and Grove-White, Elizabeth, "The Endings Project", *The Endings Project*, 2016 -. https://projectendings.github.io (accessed April 27, 2018).

Cellan-Jones and Lee 2017 Cellan-Jones, Rory and Lee, Dave, "Massive Cyber-Attack Hits 99 Countries", *BBC News*, May 13, 2017, sec. Technology. http://www.bbc.co.uk/news/technology-39901382.

Ciula 2017 Ciula, Arianna, "Call for Expressions of Interest: Your Novel Idea of Publication", *King's Digital Lab*, November 22, 2017. https://www.kdl.kcl.ac.uk/blog/call-expressions-interest-your-novel-idea-publication/ (accessed February 21, 2018).

Ciula, Nyhan, and Moulin 2013 Ciula, A, Nyhan, J, Moulin, C, "Science Policy briefing on research infrastructures in the digital humanities: landscapes, ecosystems and cultures". *Lexicon Philosophicum*, 1, p. 277-287, 2013.

Conway 2010 Conway, Paul, "Preservation in the Age of Google: Digitization, Digital Preservation, and Dilemmas." *Library Quarterly* 80, no. 1 (January 2010).

Dombrowski 2014 Dombrowski, Quinn. "What Ever Happened to Project Bamboo?" *Literary and Linguistic Computing* 29, no. 3 (September 1, 2014): 326–39.

Dourish and Bell 2014 Dourish, Paul and Bell, Genevieve, *Divining a Digital Future: Mess and Mythology in Ubiquitous Computing*. Cambridge, MA: The MIT Press, 2014.

ESF 2011 European Science Foundation (ESF), Science Policy Briefing (SPB) on Research Infrastructures in the Digital Humanities, 2011 http://archives.esf.org/fileadmin/Public_documents/Publications/spb42_RI_DigitalHumanities.pdf

Edwards et al. 2013 Edwards, P. N., Jackson, S. J., Chalmers, M. K., Bowker, G. C., Borgman, C. L., Ribes, D., Burton, M., & Calvert, S. (2013) *Knowledge Infrastructures: Intellectual Frameworks and Research Challenges*. Ann Arbor: Deep Blue. http://hdl.handle.net/2027.42/97552 (accessed February 21, 2018)

European Commission 2017 European Commission. "Sustainable European research infrastructures. A call for action. Commission staff working document: long-term sustainability of research infrastructures." October 19, 2017 http://doi.org/10.2777/76269

Eyers 2013 Eyers, Tom, "The Perils of the 'Digital Humanities': New Positivisms and the Fate of Literary Theory", *Postmodern Culture*, Volume 23 - Number 2 - January 2013, http://www.pomoculture.org/2015/07/08/the-perils-of-the-digital-humanities-new-positivisms-and-the-fate-of-literary-theory/ (accessed February 21, 2018)

Hirsch 2012 Hirsch, Brett D.. *Digital Humanities Pedagogy: Practices, Principles and Politics*. Open Book Publishers, 2012.

Jakeman 2018 Jakeman, Neil, "A Low Bandwidth Solution for Cultural Heritage Web Content", *King's Digital Lab*, January 17, 2018. https://www.kdl.kcl.ac.uk/blog/african-rock-art/ (accessed February 21, 2018).

Liu et al. 2018 Liu, Alan et al. "Critical Infrastructure Studies." MLA 2018. New York City, September 6 2018.

Montfort and Bogost 2009 Montfort, Nick and Bogost, Ian, *Racing the Beam the Atari Video Computer System*. Cambridge, MA: MIT Press, 2009.

Nicholson 2018 Nicholson, Craig. "Keeping the Lights On". *Research Europe*, February 22 (2018): 13.

Open Research Data Task Force 2017 The Open Research Data Task Force, "Research Data Infrastructures in the UK." Universities UK, 2017 http://www.universitiesuk.ac.uk/policy-and-analysis/research-policy/open-science/Pages/open-research-data-task-force.aspx#sthash.oMf7Po5F.dpbs (accessed February 21, 2018).

Ramsay and Rockwell 2012 Ramsay, Stephen and Rockwell, Geoffrey, "Developing Things: Notes toward an Epistemology of Building in the Digital Humanities", in Matthew K. Gold (ed.), *Debates in the Digital Humanities*. Minneapolis: Michigan University Press, 2012, p. 75–84.

Rockwell 2010 Rockwell, Geoffrey, "As Transparent as Infrastructure: On the Research of Cyberinfrastructure in the Humanities", *OpenStax CNS* (April 2010). http://cnx.org/contents/fd44afbb-3167-4b83-8508-4e70885b6136@2/As_Transparent_as_Infrastructu (accessed February 21, 2018).

Rosenthalter et al. 2015 Rosenthalter, Lukas, Fornaro, Peter, Immenhauser, Beat, and Böni, David, *Final report for the pilot project "Data and Service Center for the Humanities" (DaSCH)*, Swiss Academies Reports 10, 1, 2015. Zenodo. http://doi.org/10.5281/zenodo.822898

Rusbridge 2007 Rusbridge, Chris, "Arts and Humanities Data Service Decision", *Digital Curation Centre*, June 6, 2007. http://www.dcc.ac.uk/news/arts-and-humanities-data-service-decision (accessed April 27, 2018).

Russell and Vinsel 2016 Russell, Andrew and Vinsel, Lee, "Hail the Maintainers." *Aeon*. 2016. https://aeon.co/essays/innovation-is-overvalued-maintenance-often-matters-more (accessed February 21, 2018).

Sayers 2017 Sayers, Jentery (ed.), *Making Things and Drawing Boundaries*. Minneapolis: University of Minnesota Press, 2017.

Short et al. 2012 Short, Harold, Nyhan, Julianne, Welsh, Anne and Salmon, Jessica. "'Collaboration Must Be Fundamental or It's Not Going to Work': an Oral History Conversation between Harold Short and Julianne Nyhan." "Digital Humanities Quarterly" 6 no 3 (2012). Retrieved from http://www.digitalhumanities.org/dhq/vol/6/3/000133/000133.html

Smithies 2017 Smithies, James, *The Digital Humanities and the Digital Modern*. Basingstoke: Palgrave Macmillan, 2017.

Smithies 2018 Smithies, James. "Full Stack DH: Building a Virtual Research Environment on a Raspberry Pi", In *Making Things and Drawing Boundaries: Experiments in the Digital Humanities*, edited by Jentery Sayers. Debates in DH. Minneapolis: University of Minnesota Press, 2018, pp.102-114.

Stefik and Hanenberg 2014 Stefik, Andreas and Hanenberg, Stefan, "The Programming Language Wars: Questions and Responsibilities for the Programming Language Community", in *Proceedings of the 2014 ACM International Symposium on New Ideas, New Paradigms, and Reflections on Programming & Software*. New York, NY, USA: ACM, 2014.

Svensson 2016 Svensson, Patrik, *Big Digital Humanities: Imagining a Meeting Place for the Humanities and the Digital*. Ann Arbor, MI: University of Michigan Press, 2016.

Tanner 2012 Tanner, Simon, *Measuring the Impact of Digital Resources: The Balanced Value Impact Model*. King's College London, October 2012.

Turner 2008 Turner, Fred, *From Counterculture to Cyberculture: Stewart Brand, the Whole Earth Network, and the Rise of Digital Utopianism*. Chicago: University of Chicago Press, 2008.

Verhoeven 2016 Verhoeven, Deb. "As Luck Would Have It." *Feminist Media Histories* 2, no. 1 (2016): 7–28

Verhoeven and Burrows 2015 Verhoeven, Deb, and Toby Burrows. "Aggregating Cultural Heritage Data for Research Use: The Humanities Networked Infrastructure (HuNI)." in Emmanouel Garoufallou, Richard J. Hartley, and Panorea Gaitanou (eds.), *Metadata and Semantics Research*, Springer International Publishing [Communications in Computer and Information Science 544], 2015, pp. 417–23.

2019 13.1 | XML | Print Article

URL: http://www.digitalhumanities.org/dhq/vol/13/1/000411/000411.html
Last updated: January 26, 2021
Comments: dhqinfo@digitalhumanities.org
Published by: The Alliance of Digital Humanities Organizations and The Association for Computers and the Humanities
Affiliated with: Digital Scholarship in the Humanities
DHQ has been made possible in part by the National Endowment for the Humanities.

Copyright 2005 - 2021
This work is licensed under a Creative Commons Attribution-NoDerivatives 4.0 International License

NATIONAL ENDOWMENT FOR THE HUMANITIES